DOCUMENT RESUME

ED 070 716 SO 005 045

AUTHOR	Popham, W. James
TITLE	Designing Teacher Evaluation Systems. A Series of Suggestions for Establishing Teacher Assessment Procedures as Required by the Stull Bill (AB293), 1971 California Legislature.
INSTITUTION	Instructional Objectives Exchange, Los Angeles, Calif.
PUB DATE	Dec 71
NOTE	51p.
AVAILABLE FROM	The Instructional Objectives Exchange, Box 24095, Los Angeles, California 90024 ($1.25)
EDRS PRICE	MF-$0.65 HC Not Available from EDRS.
DESCRIPTORS	*Effective Teaching; Evaluation Criteria; *Evaluation Methods; Evaluation Techniques; Faculty Evaluation; Systems Analysis; Systems Approach; *Systems Development; Teacher Characteristics; *Teacher Evaluation

ABSTRACT

The analysis is intended to provide California educators with a number of specific, step-by-step suggestions for devising the teacher appraisal systems required by new legislation in the Stull Bill. An overview of essential ingredients of the new requirements is given in the first part of this document in order to familiarize teachers with required elements of the new legislation so they may consider alternative methods of implementation. Short sections deal with preliminary considerations of the analysis and a brief review of unsatisfactory methods of assessing teacher effectiveness which have been employed over the years. Major emphasis is upon designing a satisfying system of appraising teachers. Each of the key elements in the new legislation is examined in detail, and is followed by a series of possible implementation procedures. A verbatim section of the new law dealing with teacher education and a selected reference section are presented at the close of the report. (SJM)

ED 070716

DESIGNING TEACHER EVALUATION SYSTEMS

A Series of Suggestions for Establishing
Teacher Assessment Procedures
As Required by the Stull Bill (AB 293),
1971 California Legislature

SP 005 045

W. James Popham

FILMED FROM BEST AVAILABLE COPY

The Instructional Objectives Exchange
Box 24095
Los Angeles, California 90024

December, 1971

PERMISSION TO REPRODUCE THIS COPYRIGHTED MATERIAL BY MICROFICHE ONLY HAS BEEN GRANTED BY

W. James Popham

TO ERIC AND ORGANIZATIONS OPERATING UNDER AGREEMENTS WITH THE U.S. OFFICE OF EDUCATION. FURTHER REPRODUCTION OUTSIDE THE ERIC SYSTEM REQUIRES PERMISSION OF THE COPYRIGHT OWNER

PREFACE

On January 28, 1971, Assemblyman John Stull introduced AB 293 during the 1971 Regular Session of the California Legislature. Although later joined by Senator Albert Rodda and Assemblyman Leo Ryan as co-authors, the legislation is popularly referred to as the Stull Bill. Having been amended many times, it was finally enacted by the Legislature and signed by the Governor on July 20, 1971, as Statutes 1971, Chapter 361.

Because the new law stipulates that local school districts must set up their own teacher evaluation systems, there appeared to be a need for an analysis of the options available to local educators. While the Stull Bill does specify the inclusion of certain ingredients in the locally devised evaluation systems, there is still considerable room for imaginative local implementation of the new law. Accordingly, the Instructional Objectives Exchange, a California-based nonprofit corporation, is providing this booklet as a service to the California educators who must design systems to satisfy these new legal requirements. Hopefully, the analysis and suggestions contained in the following pages will be of utility to the administrators and teachers who must implement the Stull Bill.

There is some uncertainty regarding when the teacher evaluation phases of this new law become operative. Depending on the data of the final adjournment of the 1971 California Legislature, there is a possibility that the teacher evaluation requirements could become effective during the 1971-72 academic year. Certainly, they will become effective during the 1972-73 academic year. Under either circumstance, it is apparent that local educators have no time to waste in considering how to develop a defensible system of teacher appraisal.

W. James Popham
Los Angeles
December 1, 1971

TABLE OF CONTENTS

Preface..... i

Introduction..... 1

Essential Ingredients of the New Requirements -- An Overview..... 2

- Local District Adoption 2
- Teachers to be Consulted 3
- Written Report plus Face-to-Face Conference Required 3
- Probationary Teachers Evaluated Annually; Permanent Teachers Evaluated Biennially 4
- Areas of Needed Improvement to be Identified 4

Preliminary Considerations..... 5

Assessing Teacher Effectiveness -- A Brief Background Analysis..... 7

- Ratings 8
- Observations 9
- Standardized Tests 10

Designing the System..... 13

- Local District Adoption 13
- Teachers to be Consulted 16
- Certain Evaluation Elements Stipulated 19
- Assessment of Adjunct Duties 27
- Suitable Control and Learning Environment 30
- Additional Criteria 34
- Written Reports plus Face-to-Face Conferences Required 39
- Probationary Teachers Evaluated Annually; Permanent Teachers Evaluated Biennially 41
- Some Final Thoughts 42

Appendix..... 45

Selected References..... 48

INTRODUCTION

This document was prepared in direct response to the enactment of Assembly Bill 293 by the California Legislature during the 1971 regular session. This new law focused on amendments to the Education Code dealing with tenure of certificated personnel, but also required the establishment of teacher appraisal systems in each school district of the state. Article 5.5 of the law (Evaluation and Assessment of Performance of Cerfificated Personnel)* sets up some very specific requirements for these new evaluation systems, one of which is that the governing board of each district adopt "objective evaluation and assessment guidelines" after consulting the local teachers' organization. Thus, both school boards and teachers should become conversant with the required elements of the new legislation so that they can more intelligently consider alternative methods of implementing the new requirements.

The analysis in the remainder of this essay will attempt to isolate those aspects of the new legislation which mandate particular forms of appraisal systems and those which offer local options. When local choice is possible, alternative procedures for satisfying the legislation will be spelled out. The analysis and suggestions contained hereafter are designed to be practical ways of helping those who must establish the required

---

* A complete copy of Article 5.5 is included in the Appendix. See pages 45-47.

evaluation and appraisal systems. Hopefully, by presenting an expanded range of alternative techniques with which teacher evaluation systems can be designed, those responsible for setting up the systems can make more enlightened decisions regarding the system suitable for their particular school district. The tone of the document, therefore, will definitely not be prescriptive but, rather, will be suggestive of multiple options which can be put together in a defensible system of teacher appraisal.

Essential Ingredients of the New Requirements -- An Overview

It will be useful at the outset to provide a brief overview of all of the elements in the new legislation related to the teacher evaluation systems (Article 5.5). Having supplied this overall picture of the new requirements, each specific section of the legislation will be treated at greater length, along with an examination of alternative ways which might be used to satisfy the requirements. At any point the reader may find it useful to consult the Appendix which contains the verbatim section of the legislation under analysis. The following, then, are the key elements in the new law.

* * *

Local District Adoption

Each school district in the state shall develop and adopt its own objective evaluation and assessment guidelines (Section

13485).

Teachers to be Consulted

In developing a new evaluation system the governing board is to avail itself of the advice of the district's organization of certificated personnel (Section 13486).

Certain Evaluation Elements Stipulated

The evaluation system adopted must at least include the following elements (Section 13487):

(a) established standards and techniques for assessing student progress in each area of study.

(b) assessment of teacher competence as it relates to the established standards.

(c) assessment of duties performed by teachers as an adjunct to their regular assignments.

(d) established procedures for ascertaining that teachers are maintaining suitable control and learning environments.

Written Report plus Face-to-Face Conference Required

The evaluation must be transmitted in writing to the teacher at least 60 days before the close of the school year. A written response by the teacher will be included in the teacher's personnel file. A face-to-face meeting between

evaluator and teacher is required before the end of the school year (Section 13488).

Probationary Teachers Evaluated Annually: Permanent Teachers Evaluated Biennially

All probationary teachers must be evaluated at least once each school year. For teachers with permanent status, an evaluation every other year is required (Section 13489).

Areas of Needed Improvement to be Identified

The written evaluation must include recommendations, if necessary, regarding areas of the teacher's performance which should be improved. The employing authority should endeavor to assist the teacher in improving areas of performance deemed unsatisfactory (Section 13489).

* * *

Now that the chief stipulations of the new legislation have been previewed, we can turn to a more intensive analysis of each of these points, complete with series of atlernative suggestions for implementation. But before dealing with the particulars, a few preliminary observations are in order.

Preliminary Considerations

It should be made clear at the outset that the following suggestions for implementation do not imply either approval or disapproval of the particular legislative act which stimulated the analysis. The simple fact is that California educators are now obliged by their elected lawmakers to set up and operate systems of teacher appraisal and evaluation. If such systems have to be established, it seems clear that they should be designed as effectively as possible. Accordingly, this analysis attempts to put forward some options which will hopefully be blended by local teachers, administrators, and school boards into a truly defensible scheme for teacher appraisal.

Secondly, the focus of the following analysis is on <u>teacher</u> assessment and evaluation, not on the appraisal of all types of certificated personnel. It is true that Assembly Bill No. 293 (the Stull Bill) which created the evaluation and assessment provisions deals with all certificated personnel. The chief implementation difficulty, however, at least in terms of the magnitude of the enterprise, will be teacher appraisal. Hence, the current document is restricted to a consideration of teacher evaluation systems. It is likely, of course, that some of the observations made here will be pertinent to the appraisal of other types of certificated personnel.

Another consideration deals with two terms which are used often in the new legislation, namely, "evaluation" and "assessment".

For many educators these two terms are essentially interchangeable. For others, there are subtle but important differences. To illustrate, some people use the term "assessment" to express an essentially nonvaluing measurement operation, i.e., to better describe the current status of a given phenomenon by attaching a numerical description to it. But "evaluation" for these same people signifies the rendering of a value judgment regarding the merit or worth of a phenomenon. For example, a student's performance on a history achievement test might be assessed by computing the number of items answered correctly. The performance would be evaluated, however, only when someone made a judgment regarding the merit of the student's historical knowledge as reflected by the assessment. Thus, in general, assessment would precede evaluation, for evaluations could be made in a more enlightened manner if accurate assessments were available.

Now it is impossible, of course, to know what these two terms meant to the lawmakers who drafted the legislation or to those who approved it. One suspects that for many legislators the distinction described above was not considered. In the remaining analysis, therefore, we shall use terms such as evaluation, assessment, and appraisal, without a great deal of technical refinement. In general, they will refer to a determination of the degree to which the teacher is functioning competently. Since the legislation emphasizes the establishment of objective evaluation and assessment guidelines, we shall assume that the

terms at least do not refer to capricious hunches about a teacher's skill but, rather, connote a systematic scheme whereby the teacher's effectiveness can be identified. This leads us to a brief examination of previous efforts to isolate that elusive research target known as teacher effectiveness.

Assessing Teacher Effectiveness -- A Brief Background Analysis

Although this is not the place for an elaborate analysis regarding the search for adequate indicators of teacher competence, it is important to set the stage for consideration of a practical system of teacher evaluation by inspecting some of the general strategic approaches which have been employed over the years in this important research area.

Since the early 1900s, both educational researchers and educational practitioners have devoted enormous energy in efforts to devise suitable indices of teacher competence. Obviously, there are a myriad of uses to which such measures could be put. Thus, it is not surprising that so many people have devoted their attention to this quest. In general, three classes of criterion measures have been employed in teacher effectiveness inquiries, namely, ratings, systematic observations, and standardized achievement tests. Each of these types of indicators, unfortunately, has proved to be far less than satisfactory.

A brief examination of each will suggest why this has been the case.

Ratings

Perhaps the most commonly employed method of assessing a teacher's skill has been through the use of ratings of the teacher's instructional ability. Usually the ratings are supplied by administrators or supervisory personnel, but ratings have also been gathered from students, colleagues, or the teacher himself. The chief difficulty with the use of ratings in identifying teacher competency is that different raters possess markedly different ideas of what constitutes teaching skill. It is an unfortunately true assertion that many beginning teachers are judged ineffectual merely because their administrator happens to observe them teaching in a way which is at variance with the techniques he used during his "outstanding" days as a classroom teacher.

Almost everyone has an idea of how the good teacher should function, but the distressing fact is that these perceptions of good teaching vary so considerably that when raters apply their criteria (usually ill-defined and sometimes inconsistent) in formulating a rating of a teacher's skill, the result is a hodge-podge. Even when ratings are broken down into very discrete sub-dimensions, as is commonly found in many evaluation forms required for personnel files, most evidence suggests that the

separate ratings are so highly intercorrelated that it would make far more sense to employ a single, overall rating. Furthermore, these overall ratings generally do not correlate particularly well with measures of pupil growth. In summary, despite their prevalent use throughout the field of education, ratings have proved almost worthless in isolating teaching competency.

Observations

Another technique that has been employed heavily since the turn of the century, and appears to be enjoying a resurgence of interest during recent years, is the use of systematic observations of the teacher's classroom performance. Highly sophisticated observation systems have been developed by researchers during the past decade, and these techniques have been employed with considerable fervor by devotees of classroom observation techniques. The main problem with classroom observation strategies is that there is tenuous evidence at best which indicates that certain teacher behaviors are indeed related across the board to student outcomes. In other words, although we are becoming increasingly sophisticated in identifying what a teacher actually does when interacting with a classroom full of students, the fantastic diversity among teachers, pupils, and instructional intentions makes the application of observation techniques difficult to use as an index of a particular individual's teaching skill.

Standardized Tests

Finally, standardized tests have been widely employed as an index of one's teaching prowess. Characteristically, these tests are administered early in the academic year and then later in the same academic year, with results in particular subject fields attributed to the teacher most actively involved in teaching that subject matter. The reason that the use of this strategy has proved unrewarding hinges on the nature of the measurement technique employed, that is, the standardized test.

In general, standardized achievement tests have been designed to produce the kind of variation among learners which permits educators to discern the degree to which different individuals compare with one another. Standardized tests have been created so that we can identify those pupils who are better, or worse, than other pupils. The necessity to produce variation among learners has led to test construction and test improvement procedures characterized as norm-referenced measurement methods. The difficulty with norm-referenced approaches is that by their very nature they often yield tests which are the least sensitive to detecting the kind of learner growth resulting from effective instruction. An examination of recent literature* in the field of criterion-referenced measurement (an alternative to norm-referenced approaches) will indicate why this is so.

---

* See selected references section for such citations.

A second difficulty with these standardized tests is that since they are all often so global in nature, it is extremely difficult to tell precisely what has been measured. Student performance reflected by a general index such as "comprehension of historical knowledge" does little to guide us regarding what a particular teacher has actually been accomplishing with learners.

Finally, because the scope of the content covered by the standardized tests is essentially a given, and because many teachers vary in their particular content emphases, the standardized test often does not mesh with the teacher's instructional preferences. As a consequence, the use of standardized tests has proved less than satisfying for assessing teacher competence, even though our familiarity with such measures through the years has made both teachers and the public comfortable with them. Nevertheless, according to many measurement experts, standardized tests are not satisfactory tools for assessing the effects of instruction.

In review, then, the three methods of assessing teacher effectiveness which have been most commonly employed during the past seventy years, that is, ratings, systematic observations, and standardized achievement tests of learner progress, have all proved deficient in one way or another. It is not possible for us today to capitalize on a well established and functional technology of teacher assessment, a technology which has been honed to a high level of precision through the years. Instead, the distressing truth is that systems for assessment and evaluation of

teacher performance must be erected on technical foundations which more closely approximate balsa wood than concrete.

Hence, having recognized at the outset that those who are commissioned to develop systematic procedures for appraising teacher success can not capitalize on a series of already proven measurement schemes, it is reasonable to assume that the resulting systems will undoubtedly be less than perfect. Yet, by calling on the best current knowledge regarding the appraisal of instructional skill we can devise systems consonant with recent advances, thereby avoiding at least serious errors of the past.

Some may question the wisdom of trying to establish procedures for evaluating teachers when the supporting assessment technology is less than adequate. This objection might have more cogency if teachers were not already being evaluated. And the current evaluations are just as influential as any future evaluations will be. Tenure decisions, assignment decisions, advancement decisions -- all of these and more are <u>currently</u> being made at least in part on the basis of teacher appraisals. Since evaluations are already being used, why not make them as defensible as possible? Moreover, since the ultimate reason for evaluating teachers' instructional skills is to provide a better education for students, the goal of improved teacher evaluation seems well worth the effort.

Designing the System

Now we can turn to the task of designing a satisfactory system for appraising teachers. Each of the key elements in the new legislation will be examined in some detail, followed by a series of possible implementation procedures.

Local District Adoption

We can start by considering the major purpose of the new law insofar as it bears on teacher appraisal. This purpose is clearly set forth in the legislation:

> 13485. It is the intent of the Legislature to establish a uniform system of evaluation and assessment of the performance of certificated personnel within each school district of the state. The system shall involve the development and adoption by each school district of objective evaluation and assessment guidelines.

There are several important points in this statement of intent. First, notice that the legislature wishes to establish a uniform system of teacher appraisal within each district. This means, quite clearly, that for districts involving several schools or more it is anticipated that district-wide guidelines will be established. It would appear that the evaluation system is to be consistent throughout the district, not highly variable depending on whether School A or School B is involved. Although, at first inspection, this is a likely interpretation of the legislature's intent, further consideration will reveal that a uniform system can be responsive to differences within the

district. For example, a mechanism could be uniformly installed in the evaluation system whereby particular schools could devise additional inputs to the evaluation data. Not that this need be encouraged, for some districts might decide very judiciously to devise an essentially invariant evaluation scheme to be employed across the district, irrespective of variations within the district. The point is that the term uniform need not preclude the possibility of certain variations within a given district.

A second point hinges on the explicit intention of setting up evaluation and assessment guidelines within each school district. It is apparent that no state-wide system of teacher appraisal is being proposed here. On the contrary, each district is to develop and adopt its own system.

A third key element of the legislature's intention revolves around the phrase "objective evaluation and assessment". If we are to interpret the term objective in the customary fashion, then the teacher appraisal system to be adopted should be as well defined and well understood as possible. No evaluation system is objective if it is primarily dependent on intuitively derived assessments. A school administrator who visits a teacher's class on one or two occasions during the year, then bolsters such impressionistic data with personal perceptions of the teacher's cooperativeness and general intelligence, could hardly be said to have engaged in an objective evaluation of the teacher. This new law requires that the ground rules be spelled out in

considerable detail so that, within the limitations of our current measurement technology, we can appraise teachers in as precise a fashion as possible.

In this connection a comparison from the educational measurement field may be illuminating. Teachers often distinguish between objective tests, such as multiple choice or true-false examinations, and subjective tests, such as essay examinations. Now the critical difference between these two approaches to testing is that for objective tests <u>different individuals</u> can score a student's response and come up with essentially the same result. Consistency of interpretation is the key. An essay test will usually be graded very differently according to who is doing the grading. The more objectivity we bring to our testing operations, the more reliably we can judge learner performance. In the same way, the new legislation requires the development of teacher evaluation systems which can be supervised by different individuals so that consistent results are yielded. No hidden or capricious criteria are to be employed. No building principal's "feeling" for what constitutes good teaching is to be used. The evaluation system is, by law, to be <u>objective</u>.

<u>Possible Implementation Procedures</u>

This is not a particularly complicated part of the new law. The chief consideration is that someone within the district must take the initiative to start devising the evaluation system.

Since this part of the law becomes effective either during the 1971-72 or the 1972-73 school year, its requirements are mandated within a few months. Hence, there is no time to waste in getting underway. The individual (or individuals) responsible for developing the system should probably do some quick background reading regarding teacher effectiveness assessment from such sources and those cited in the selected references section (pages 48-49).

Since, as will be seen in later paragraphs, teachers, administrators, and board members will ultimately be involved in deliberations regarding the new instructor evaluation system, it would seem prudent to get representatives of such groups engaged in preliminary explorations immediately.

Of course, an alternative course of action is to take no action for the time being with the hope that more definitive suggestions will emerge from groups such as the various California professional associations whose members are affected by the legislation, e.g., school boards, school administrators,* teachers, and teacher educators. The only problem with this temporizing strategy is that if the suggestions produced by these groups are not of sufficient quality or specificity, then valuable planning time will have been irretrievably lost.

Teachers to be Consulted

A critical procedural feature of the new law requires that

---

* E.g., see the excellent Management Action Paper, Association of California School Administrators by Thomas Shannon, October, 1971.

in creating the guidelines for the teacher appraisal system "the governing board shall avail itself of the advice of the certificated instructional personnel in the district's organization of certificated personnel." A school board which adopted a teacher evaluation system without consulting the district's teachers would clearly be in violation of the legislation.

Brief consideration of the legal requirement contained in the wording of the legislation makes it evident that a school board would satisfy the letter of the law with a one-shot request to the local teaching organization for their thoughts, then completely disregard such suggestions. Yet, it seems unlikely that a school board composed of reasonable citizens would adopt such a tokenistic strategy. For since teachers will be vitally affected by the new system, failure to involve their representatives in the development of the appraisal system would surely engender a continuing source of irritation and uncooperativeness.

By the same logic, teacher groups who have been invited to participate seriously in devising a set of appraisal guidelines would be doing a disservice to themselves by not approaching the task as constructively as possible. While teachers' employment interests must undoubtedly be protected by their representatives, a constructive spirit will undoubtedly lead to development of a more satisfactory evaluation system.

There may be situations, particularly in smaller districts, where there is not an officially designated teachers' organization.

In such cases teacher representatives can surely be identified, possibly through informal elections. In instances where two or more major organizations represent the district's teachers, representatives from all such groups should undoubtedly be involved.

Possible Implementation Procedures

There are numerous possibilities whereby a school board may avail itself of the advice of the district's teachers. As suggested earlier, the spirit of the consultation with teachers may range from tokenistic to total. At the more comprehensive level, the board or its administrator-representatives can involve teachers at the very outset in deliberations regarding the design of an evaluation system. For example, a district assistant superintendent might be commissioned to develop the appraisal system working with an advisory group constituted by such individuals as teachers, other administrators, perhaps a board member or two, possibly a specialist in evaluation from a nearby university, and maybe a few students and parents. If the advisory committee is really influential in designing the system for board adoption, then the teacher representatives would clearly have an opportunity to express their colleagues' concerns.

All sorts of variants are possible for consulting with teachers. For instance, those given the responsibility for designing

the system might submit all preliminary drafts describing the system to a representative teachers' group for their reactions. It might even be possible to ask a teachers' group to generate an initial draft of the proposed evaluation system.

The really imperative consideration here is that both the board and the district teachers recognize the necessity to participate in developing the new teacher assessment system. Representatives of the local teachers' group should contact the board's representatives as soon as possible to discuss alternative ways of incorporating the requisite participation of teachers.

Certain Evaluation Elements Stipulated

So far in our discussion we have examined those aspects of the new law which are not relatively constraining regarding procedures for setting up a teacher appraisal system. Now we turn to four stipulations which require those designing the system to incorporate particular criteria. In many ways the inclusion of these required elements represent the really important feature of this new legislation, for there are important criteria included which are often absent from current teacher evaluation systems. We shall consider each of the four required elements separately.

It should be noted, however, that these four elements are minimum criteria, and that other considerations can be built

into the system. This is set forth quite clearly in the new legislation:

> 13487. The governing board of each school district shall develop and adopt specific evaluation and assessment guidelines which shall include but shall not necessarily be limited in content to the following elements: . . .

We can turn now to an inspection of the first of the four required elements, examining additional criteria at a later point.

Establishing standards of expected student progress and assessment techniques. This particular requirement represents perhaps the most significant feature of the new law, for it establishes quite clearly that one important criterion in the new teacher evaluation system must be based on student growth. Because of its importance, we should consider the language of the legislation carefully:

> (a) The establishment of standards of expected student progress in each area of study and of techniques for the assessment of that progress.

The phrase "establishment of standards of expected student progress" would seem readily interpretable. A local school district will have to decide upon some well defined levels of student performance which are considered acceptable for that district. Use of the term progress implies a before-after conception of student performance and, thus, is probably equivalent to the familiar notion of learner growth often tapped through the use of pretest-posttest designs involving the assessment of an instructional procedure. How to establish progress standards

which will be considered acceptable is, of course, a sticky question, but we can consider some alternative ways of going about this in a moment.

Another phrase which seems pretty straightforward indicates that these student progress standards must be established "in each area of study." Here, one suspects, the legislature was asking for standards in each of the subjects taught in our schools. Hence, in the secondary schools standards of student progress are anticipated for each course taught, e.g., history, English, or biology. For the elementary school we would expect standards to be established for all of the major curricular emphases, e.g., reading, mathematics, and language arts.

Finally, the phrase which dictates the establishment of "techniques for the establishment of that progress" makes it clear this law anticipates that more than local rhetoric will be produced. Instead of lofty proclamations about having "our children attain outstanding levels of intellectual excellence," local educators will have to explicate assessment procedures which permit the unambiguous determination of whether the expected standards have been achieved.

Possible Implementation Procedures. There are several ways available for getting at the standards of student progress problem. If we interpret the progress construct literally, then there will clearly have to be more than one assessment made up over an extended temporal period, such as at the beginning

and close of a school year. One of the most obvious alternatives which will occur to many is to employ standardized achievement tests on a pretest-posttest basis by administering such measures each Spring. By using each such measurement as a posttest for the academic year in which it is administered and a pretest for the ensuing academic year, the district can get by with one major testing operation each year.

There are severe problems with standardized tests which should be recognized. First, as mentioned earlier in this discussion, most standardized tests have been devised with a view to assessing learners' status in relationship to one another, not to assessing pre- to post-instruction growth. In certain respects some achievement tests are too insensitive to the detection of student progress to suit our purposes. Second, since tests will have to be employed in each area of study, the costs of acquiring commercially produced standardized tests might be prohibitive. Third, because we probably do not wish to take away an enormous amount of instructional time for purposes of testing, we may wish to employ person sampling and item sampling techniques. Item sampling involves different individuals completing different segments of a test, that is, each pupil taking less than the total test, in order to yield an overall estimate of group progress. Many standardized tests, because they are copyrighted, do not lend themselves readily to dismembering and local reproduction for item sampling purposes.

An alternative to standardized tests is to employ criterion-referenced examinations which are now being distributed by an increasing number of agencies. Some of these new tests are being distributed intact.* Others require the local districts to assemble them based on the preferences of district educators regarding what elements should go into the test.** In either instance, local school people can select tests or test items without having to start from scratch in developing their own measures.

Another option, of course, is to have local educators construct new measuring instruments to be employed in assessing the expected standards of student progress. Because of the comprehensiveness of the assessments to be undertaken, of course, this will be an expensive enterprise.

In any event, measures of learner performance in all subject areas must be developed, adopted, or adapted. To secure a defensible estimate of the progress of a teacher's pupils it has been suggested that item sampling be combined with pupil sampling to consume less testing time. The period of instruction during

---

* See, for example, the measures being distributed by Educational and Industrial Testing Service, P.O. Box 7234, San Diego, California 92107.

** For example, objectives-bank agencies and test item-bank agencies such as the Instructional Objectives Exchange (Box 24095, Los Angeles, California 90024) distribute a wide array of objectives and related criterion-referenced test items which can be readily assembled by local educators for use in item sampling or non-item sampling examinations.

which the expected pupil progress is to manifest itself has not been determined. Although many educators will automatically think of whole year periods, there may be some practical advantages to opting for much shorter periods.

Now that we have considered a number of schemes for securing measuring instruments, we still have not resolved the problem of how to establish the standards of expected progress. There are a number of ways to get a fix on such standards. We recognize immediately that the establishment of such standards is a value laden decision on the part of whoever is responsible for establishing the standards. Nevertheless, someone has to establish them. Considering the fact that parents, teachers, and even the learners themselves, have a stake in determining what the standards should be, we might consider various ways of including such groups.* For example, once the measuring instruments have been determined, then the district might consult with (a) parents, (b) teachers, and (c) non-parent community representatives to determine what percentage of the district's children should display what degree of proficiency on the measures. Perhaps these deliberations can be made more meaningful if a sample of current learner performance on the measures can

---

* For descriptions of systematic schemes for involving diverse clienteles in such decision-making, see some of the recent materials dealing with educational needs assessment, e.g., Determining Instructional Goals via Educational Needs Assessment, a filmstrip-tape program distributed by Vimcet Associates, P.O. Box 24714, Los Angeles, California 90024.

be secured for consideration. Although status-quoism must be guarded against, there are clear dividends to be gained from knowing roughly how learners are presently performing.

Perhaps each subject area teacher group, or a representative group if the district is too large, can consider how to select measuring instruments, then with or without others' involvement, can set some tentative standards for their area of concern. In this endeavor they should recognize that comprehensiveness of coverage is not dictated by the law. As long as one or more standards of expected student progress are spelled out for each area of study, the law is satisfied. Thus, if a district's resources do not permit development of a complete measurement scheme in which all or most important areas of anticipated student growth were assessed, it would be sufficient to identify only one or two important educational goals for each area of study and assess learner progress toward these goals. Obviously, the more comprehensive our system, the more validly it can be used to reflect the instructional quality of a teacher. Yet, practical constraints will sometimes force local educators to select a less than totally comprehensive assessment scheme. The important first step is to establish some expectations, and spell them out in clear (preferably quantifiable) terms. The comprehensiveness of the system can be expanded at a later point.

Relating teacher competence to the establishment standards.

We can turn now to the second required element in the teacher evaluation system, since it hinges directly on the first, that is, the established standards of pupil progress. Quite clearly the legislature anticipates a direct effort to tie down the degree to which a given instructor is contributing toward his students' progress (or lack of it). The exact wording of this second stipulated element reads as follows: "Assessment of certificated personnel competence as it relates to the established standards."

Possible Implementation Procedures. On this point there do not seem to be too many options. One clear alternative is to make certain that a reasonable sample of each teacher's pupils are involved in any assessment operations so that pre- and post-instruction indices can be calculated for each teacher. When enough students are available, random selection techniques should be used to reduce inadvertent bias in sample selection.

Possibly the appraisal system could be designed so that the teacher would play a prominent role in gathering and reporting the data. Yet, in the interest of objectivity it would appear that an external data gatherer, e.g., an administrator or impartial colleague, should be the chief data collector.

A routine reporting form would undoubtedly have to be devised in which the growth of each teacher's pupils toward the established standards could be displayed so that both the teacher and other involved parties could readily comprehend the implications

of the data.

Assessment of Adjunct Duties

We turn now to a consideration of the third of the four required elements stipulated in the new law for inclusion in the district-developed teacher appraisal system. An examination of the verbatim wording of this section will be instructive: "Assessment of other duties normally required to be performed by certificated employees as an adjunct to their regular assignments." There would appear to be a possible contradiction involved in this particular phrasing, for to the extent that duties are "normally required" it seems that they would also be considered a part of a teacher's "regular requirements". Yet, it is more likely that no subtle semantic distinctions are involved in this third required element, but instead the reference is to non-teaching duties. For many legislators, a teacher's primary responsibility is discharged via classroom instruction, thus non-instructional responsibilities are probably those "other duties" referred to in this segment of the legislation. At any rate, this is the interpretation which guides the subsequent discussion.

Possible Implementation Procedures

We have a very difficult problem here, for on the one hand the local district is urged to establish an objective evaluation

system and yet we now encounter a criterion which would appear to involve someone's judgment regarding the degree to which a teacher is successfully performing such adjunct duties as the supervision of extracurricular activities, participation in faculty curriculum committees, etc. Such "judgment calls", whether rendered by supervisor, administrator, or other teachers, are notoriously subjective rather than objective and put the designers of a district's teacher appraisal system in a difficult position.

Perhaps the only thing that can be done in this situation is to make as systematic as possible the procedures which must be used for reaching judgments regarding the proficiency with which a teacher discharges non-instructional responsibilities. Since a chief consideration in evaluating the degree to which a teacher carries out non-instructional responsibilities is the actual participation of the teacher in the activity (as opoosed to the quality of that participation), it is probably conceivable to devise some type of rating form which hinges very heavily on the teacher's participatory activities. For instance, if extracurricular activities are considered among those suitable for this particular phase of the evaluation system, then one could include participation in extracurricular activities as one factor to be judged on the basis of degree of the teacher's participation in such endeavors. For instance, one could use a clock hour participation scheme in which the approximate number of hours per academic year invested in the

activity were identified. Or, again, a scheme could be used in which a rating of the degree of participation was based on some simple category system such as adequate, superior, etc. In addition to quantitative indicators, of course, one could also set up some kind of quality judgment regarding the nature of the teacher's participation in the non-instructional adjunct activities.

One possible variation on this theme would be to have an array of adjunct activities displayed on a tentative rating form and then, possibly with teacher participation involved, assign differential degrees of import to the various activities. For instance, sponsorship of a school club might be considered important, while routine role-taking or serving as a chaperone for the school dance might be considered less important. What is being suggested is that a teacher appraisal system dealing with this particular dimension might at least try to bring some differential estimates of the import of non-instructional adjunct activities.

Another option to consider would involve the teacher's personal preparation of an itemized statement regarding the nature of the non-instructional activities participated in during the year. To do this with any degree of consistency, however, the teachers would surely have to employ some guidelines regarding how to write up their descriptions of adjunct activity participation.

Another alternative might be to rely upon colleague estimates of the degree and quality of teacher participation in such adjunct activities. Very often, the most accurate analyses regarding a faculty member's participation in such adjunct activities can be gained from the teacher's colleagues. Perhaps some simple rating form could be devised whereby teachers could rate each other, anonymously if possible, regarding the degree to which their adjunct activities had been satisfactorily performed. If such a scheme is used, it ought to be developed so that no effort is made to secure a "normal curve" where some of the teachers are, of necessity, judged either particularly good or particularly poor merely because the system is set up on a ranking rather than a rating basis. In other words, there undoubtedly would be some built-in divisiveness if colleague rating schemes were employed which, as a function of the particular appraisal mechanism utilized, automatically relegated some faculty to the receipt of an unsatisfactory judgment on this criterion.

Suitable Control and Learning Environment

This is a particularly perplexing facet of the new legislation and, because it is the fourth required element in any locally devised teacher appraisal system, we must consider it carefully. The exact working of the legislation follows: "The establishment of procedures and techniques for ascertaining

that the certificated employee is maintaining proper control and is preserving a suitable learning environment."

The problem with this particular item is that unless we wished to use it as a vehicle for identifying only those extreme instances in which a teacher is maintaining an obviously improper classroom control or is promoting an extremely unsuitable learning environment, it is next to impossible to implement. Let's see why.

The distressing fact is that unless we wish to get into a process-focused stance where we assume that by watching a teacher's classroom behavior we can tell something about what really constitutes a "suitable learning environment" or what constitutes "proper control", we are forced to turn to research evidence regarding the demonstrable elements of good control and appropriate learning environment. But there is insufficient evidence to allow us to discern with certainty what really represents an appropriate environment. The truth is that different teachers can employ markedly different techniques, yet achieve identical results. Some teachers employ very loose classroom control procedures, yet their students learn much and appear to enjoy it. Other teachers use far more stringent disciplinary tactics, yet their students learn much and seem to enjoy it. We just don't know how to isolate the constituent elements of appropriate control procedures and effective learning environment strategies. So where does that leave us with respect

to the Stull Bill?

Possible Implementation Procedures

There are a couple of strategies which might be employed with respect to the new law. One would be to assert frankly that, since there is no unequivocal evidence regarding what constitutes proper control or suitable learning environment, such dimensions must be inferred from results of student performance. And since student performance has already been required in connection with an earlier stipulation of the law, then we can infer the suitability of learning environment and the appropriateness of classroom control from the type of learner performance the teacher is able to produce. In view of the paucity of defensible research evidence regarding these considerations, such a strategy would seem eminently reasonable.

An alternative scheme which has to be considered involves the use of classroom observation activity in which a supervisor, administrator, or even some colleagues observe the teacher's classroom performance and subsequently render judgments on some kind of observation form, hopefully simple, regarding the degree to which there appeared to be gross violations of proper classroom control and suitable learning environment. Undoubtedly this would satisfy the expectations of the legislators. Yet, candidly, the stipulation of the legislature is at variance with the best that is known regarding teacher

competence assessment. All efforts to identify the critical features of a "suitable learning environment" or the necessary ingredients in "maintaining proper control" have proved fruitless. It is to be expected that legislators would not be familiar with this lack of evidence, and some might have guessed that we should have such evidence. Nevertheless, this does not alter the fact that the use of process-focused assessment techniques runs counter to the general thrust of the first two elements required by the new legislation, namely, an emphasis on demonstrable student progress.

Another option might involve having the students supply some kind of ratings of the teacher with respect to dimensions regarding control and learning environment. Student ratings are somewhat suspect for this purpose but, once more, they may well be better than nothing.

In a sense, this last consideration of how one might go about assessing the degree to which a teacher preserves an appropriate classroom environment leads to a consideration of additional criteria which might be considered by those developing the new evaluation system.

Additional Criteria

As those individuals designing the teacher evaluation system for each district become more conversant with literature in this field, they will undoubtedly consider other criteria than

either those required by the law or to be discussed in the next few paragraphs. Nevertheless, perhaps an examination of some alternative criteria will be of utility.

Use of Affective Measures. As noted previously, the emphasis of the new legislation, insofar as it relates to the assessment of pupil progress, deals with progress in "each area of study". But the law does not touch at all upon those important aspects of student growth which are more general, namely, the kinds of affective considerations which might reasonably be expected to emerge as a consequence of an instructional program. Affective considerations such as the attitudinal, interest, and valuing behaviors of learners are, in the eyes of many educators, equally if not more important than the kinds of cognitive outcomes which are to be assessed in connection with the new legislation.

Until recently, however, educators' efforts to assess the elusive affective consequences of instruction remained largely at a rhetorical level rather than at an actual measurement level. In the past several years, however, an increasing number of educational measurement specialists have turned their attention to the construction of criterion-referenced measurement techniques which can be employed to assess important attitudinal outcomes of education. It is currently recommended that such indicators be carefully surveyed for possible inclusion in

the teacher appraisal system.* Those affective measures considered suitable could be readily employed at the local level. It is also possible, of course, to construct new affective indicators within the local district, but generally this proves too time-consuming to be a realistic option. On the other hand, some very rudimentary anonymous self-report devices can be generated which allow the pupils to rate teachers in terms of such factors as the interest they can promote in the subject matter, etc.

Teaching Performance Tests. Another possible addition to the teacher appraisal system, and one which many believe represents valuable augmentation of our appraisal techniques, is the use of a teaching performance test or, as it is sometimes called, an instructional minilesson. In brief, this approach takes cognizance of the fact that it is next to impossible to compare different teachers who are not pursuing similar instructional emphases. In other words, if one U.S. history teacher emphasizes a certain aspect of the Industrial Revolution and another U.S. history teacher emphasizes the Civil War, then it is really illegitimate to contrast their performance insofar as learner achievement is concerned, for

---

* For example, the Instructional Objectives Exchange (P.O. Box 24095, Los Angeles, California 90024) now distributes collections of instructional objectives and measuring instruments dealing with students' attitude toward school, self-concept, and tolerance.

both of the teachers have been pursuing different goals. A teaching performance test overcomes this difficulty by providing an identical instructional objective for different teachers and then contrasting their ability to achieve these identical objectives as reflected by learner post-instructional performance.

Briefly, here is how a teaching performance test works. An instructor is given an explicit instructional objective, specified in terms of measurable learner behavior, along with a sample test item to help reveal exactly how the objective is to be assessed. In addition, background information is given to the teacher if the topic is unfamiliar in order that the teacher can plan a lesson for a short period of instructional time, as brief as 15 minutes. The teacher is then given a certain amount of planning time, perhaps an hour or two, during which an instructional plan is devised to accomplish the prespecified instructional objective. Next, a small group of learners (6-8 students), randomly selected from a pool of appropriate learners, is instructed by the teacher. After the instruction a posttest, not previously seen by the teacher but readily inferrable from the instructional objective and sample test item, is administered to the students. The pupils are also asked to supply an affective rating of the instruction, such as the degree to which they found the topic interesting. The performance of the students on the posttest and their affective ratings of the instruction serve as an indication

of the degree to which the teacher is skilled at this particular task, namely, the accomplishment of pre-specified instructional objectives with positive learner affect.

There are significant advantages in using short term performance tests in teaching appraisal systems. First, because of their brevity, they can be practically administered in connection with a large scale teacher appraisal system. There is at least one agency* which has already begun to offer appraisal services involving such performance tests.

Second, because identical objectives are used, it is possible to secure rather precise comparisons among teachers in a given district. Using other approaches, such relative rankings of instructional proficiency are rare. Reasonably accurate contrasts between teachers can be made in view of the fact that because randomly assigned groups of learners are used, entry skill differences between pupils are minimized. Beyond those which are controlled through random assignment, additional statistical compensation techniques have been devised which permit more careful contrasts among teachers with respect to this particular instructional skill.

---

* Instructional Appraisal Services, a newly established firm, anticipates offering services to districts in which performance tests will be conducted along the lines indicated in this document. For further information, contact Instructional Appraisal Services at 105 Christopher Circle, Ithaca, New York 14850, or Box 24821, Los Angeles, California 90024.

It should be emphasized, however, that the ability to accomplish pre-specified objectives in learners under short instructional periods represents only one criterion by which a teacher should be judged. The use of teacher performance tests simply reflects an additional criterion which might be employed in a local district's evaluation system.

Incidentally, as we shall see in a later part of the legislation, local administrators are required to confer with employees judged deficient and make recommendations as to areas of improvement in the employee's performance. The use of practice self improvement kits whereby teachers can attempt to improve their skills on this particular instructional task are also beginning to be produced.* Short term minilessons conducted by teachers, followed by critique sessions involving other teachers and, possibly, supervisors auger well for the improvement of this aspect of a teacher's instructional proficiency.

Beyond the four required elements and the inclusion of additional criteria, there are remaining procedural stipulations in the new legislation. Most of these are relatively self-explanatory, but they should be considered in order to complete our examination of the new law.

Written Reports plus Face-to-Face Conferences Required

The evaluation emerging from the teacher assessment system

---

* Instructional Appraisal Service (Box 24821, Los Angeles, California 90024) will distribute such improvement kits commencing spring 1972

is to be prepared in written form and given to the teacher not later than 60 days before the close of the school year in which the evaluation takes place. The teacher, after receiving the evaluation, then has the right to prepare a written response to the evaluation and have such a response included as a permanent attachment to the teacher's personnel file.

In addition, before the end of the school year, a face-to-face conference is to be held between the teacher and the evaluator in order to discuss the evaluation and, possibly, the response to it. Now the evaluator(s), of course, must be identified in the teacher appraisal system. In many instances it will be a building principal. In other districts it might be a specialist assigned to this particular operation. Whoever the evaluator is, it is clear that any evaluation has to be set down in written form and forwarded to the teacher prior to the stipulated deadline. After a reasonable period in which the teacher has an opportunity to reply, a face-to-face meeting between the evaluator and the teacher should be set up. If there has been a written response from the teacher, it would be desirable to have access to the response at that time. The actual procedures associated with this written documentation and the face-to-face conference will vary from district to district.

Probationary Teachers Evaluated Annually; Permanent Teachers

Evaluated Biannually

The new legislation stipulates that probationary teachers be evaluated every year and permanent employees once every other year. It is further stipulated that the evaluation shall include recommendations, if they are necessary, regarding areas of needed improvement in the performance of the teacher. More directly, if the teacher is not performing his duties satisfactorily according to the instructional appraisal system devised by the district, it is the responsibility of the employing authority to notify the teacher in writing of that fact and to describe the areas of deficient performance. The new law also stipulates that the employing authority shall, after the written notification, confer with the teacher regarding specific recommendations as to areas of improvement in the teacher's performance and shall endeavor to aid him in the improvement of his performance.

It seems likely that in order to conserve energy, most appraisal systems will probably coalesce the two written notifications into a single effort. More specifically, the written evaluation which must be communicated to the teacher at least 60 days before the end of the academic year will undoubtedly also include suggestions for areas of improvement if such areas have been identified. In addition, it would seem that the occasion of a face-to-face meeting between evaluator and teacher would prove an ideal opportunity for suggestions regarding

methods of improving the teacher's performance.

It would seem apparent that if this particular provision of the legislation is to be satisfactorily implemented, it is necessary for the designers of the teacher evaluation system to give some consideration to the kind of tangible suggestions for improvement which can be forthcoming. This is one of the reasons why the previous identification of teaching performance tests, and the instructional minilessons which can be used for improvement, seemed like a reasonable option in this connection. Other procedures, however, should be seriously investigated so that when the evaluator and teacher actually confer, the evaluator will have better counsel than simply "Go back into the classroom and do better."

Some Final Thoughts

Now that we have considered all aspects of the Stull Bill's teacher evaluation requirements, a few final considerations are in order. First, the local option feature of the law makes it eminently clear that districts with unique programs, e.g., nongraded systems, will have to devise unique teacher appraisal mechanisms.

There are particularly vexing problems involving the evaluation of instructors who are engaged in team teaching. If the responsibilities of the team members are sufficiently discrete, then there is less difficulty. If there is truly shared responsibility, then the individual members of the team may have to be evaluated

according to the quality of the whole team. Such procedural dilemmas in atypical educational situations will have to be anticipated.

Some districts will consider the possibility of setting up highly individualized systems of teacher appraisal where an individual instructor and the evaluator set down in measurable terms the expected accomplishments of learners during the year. Such personalized "contract" systems of evaluation have been growing in popularity during recent years. The essentials of such schemes are well set forth in a recent book by McNeil.*

The overriding fear of many educators is that in their haste to set up some form of local evaluation system, California educators may turn to the omnipresent standardized achievement test. For a number of reasons cited earlier, such a decision would not only yield inaccurate information, it would probably be the scheme most penalizing to teachers, for using such measures they have the worst opportunity to display their true instructional skills.

For elementary teachers who operate in a self-contained classroom it may be that the local board will have to decide which areas of student progress will be involved in the appraisal system. It is inconceivable to imagine certain elementary teachers being evaluated on all content they cover, for that could surely range from the earthworm to the galaxies. Priorities will probably have to be set by the board.

It should be anticipated that the initial system for evaluating teachers will be flawed. But by building in a clear commitment to annually evaluate the system itself and, if necessary, revise it, then local educators can devise a self-correcting appraisal scheme which will do justice to all.

Fortunately, there seems to be a good deal of energy currently being expended to aid educators in setting up their new evaluation systems. For example, the California State Department of Education will soon be releasing suggested guidelines designed to assist in the development of local teacher evaluation systems.

In summary, the foregoing analysis was intended to provide the implementers of AB 293 with a number of specific, step-by-step suggestions for devising the teacher appraisal systems required by the new legislation. The reader is urged to consult the selected references presented at the close of the report, not to mention the verbatim section of the new law dealing with teacher evaluation (included in the Appendix).

APPENDIX

Article 5.5 Evaluation and Assessment of Performance of Certificated Employees

13485. It is the intent of the Legislature to establish a uniform system of evaluation and assessment of the performance of certificated personnel within each school district of the state. The system shall involve the development and adoption by each school district of objective evaluation and assessment guidelines.

13486. In the development and adoption of these guidelines and procedures, the governing board shall avail itself of the advice of the certificated instructional personnel in the district's organization of certificated personnel.

13487. The governing board of each school district shall develop and adopt specific evaluation and assessment guidelines which shall include but shall not necessarily be limited in content to the following elements:

(a) The establishment of standards of expected student progress in each area of study and of techniques for the assessment of that progress.

(b) Assessment of certificated personnel competence as it relates to the established standards.

(c) Assessment of other duties normally required to be

performed by certificated employees as an adjunct to their regular assignments.

(d) The establishment of procedures and techniques for ascertaining that the certificated employee is maintaining proper control and is preserving a suitable learning environment.

13488. Evaluation and assessment made pursuant to this article shall be reduced to writing and a copy thereof shall be transmitted to the certificated employee not later than 60 days before the end of each school year in which the evaluation takes place. The certificated employee shall have the right to initiate a written reaction or response to the evaluation. Such response shall become a permanent attachment to the employee's personnel file. Before the end of the school year, a meeting shall be held between the certificated personnel and the evaluator to discuss the evaluation.

13489. Evaluation and assessment of the performance of each certificated employee shall be made on a continuing basis, at least once each school year for probationary personnel, and at least every other year for personnel with permanent status. The evaluation shall include recommendations, if necessary, as to areas of improvement in the performance of the employee. In the event an employee is not performing his duties in a satisfactory manner according to the standards prescribed by the governing board, the employing authority shall notify

the employee in writing of such fact and describe such unsatisfactory performance. The employing authority shall thereafter confer with the employee making specific recommendations as to areas of improvement in the employee's performance and endeavor to assist him in such performance.

SELECTED REFERENCES

American Educational Research Association, Committee on the Criteria of Teacher Effectiveness. Report of the committee on the criteria of teacher effectiveness. Review of Educational Research, 1952, 22, 238-263.

Biddle, B. J. & Ellena, W. J. (Ed.), Contemporary research on teacher effectiveness. New York: Holt, Rinehart & Winston, 1964.

Flanders, N.A. & Simon, A. Teacher effectiveness. Encyclopedia of Educational Research, 1969, 1423-1437.

Glaser, R. Instructional technology and the measurement of learning outcomes: Some questions. American Psychologist, 1963, 18, 519-521.

Glaser, R. and Cox, R. C. Criterion-referenced testing for the measurement of educational outcomes. Instructional Process and Media Integration. (Ed.) Robert Weisgerber, Rand-McNally, Chicago, 1967.

McNeil, J. D. Toward accountable teachers: Their appraisal and improvement. New York: Holt, Rinehart and Winston, 1971.

Morsh, J. E. & Wilder, E. W. Identifying the effective Instructor: A review of the quantitative studies 1900-1952. Project No. 7714. Task No. 77243. Chenute Air Force Base, Illinois: Air Force Personnel and Training Research Center, 1953.

Musella, D. Improving teacher evaluation. The Journal of Teacher Education, Spring 1970, 21 (1), 15-21.

National Education Association, Research D vision and the American Association of School Administrators. Evaluating teaching performance. Educational Research Service Circular 3, 1969. Washington, D.C.: The Service, May 1969.

Popham, W. J. (Ed.) Criterion-Referenced Measurement: an introduction. Englewood Cliffs: Educational Technology Publications, 1971.

Popham, W. J. Performance tests of teaching proficiency: Rationale, development, and validation. American Educational Research Journal, January 1971, 8 (1), 105-117.

Rosenshine, B. & Furst, N. Research on teacher performance criteria. Teacher Training: A Symposium. (Edited by B. O. Smith), Englewood Cliffs, N.J.: Prentice Hall, Inc. 1970, 37-73.

Ryans, D. G. Characteristics of teachers: Their description comparison and appraisal. Washington, D.C.: American Council on Education, 1960.

Saadeh, I. W. Teacher effectiveness or classroom efficiency: A new direction in the evaluation of teaching. The Journal of Teacher Education, Spring 1971, 21 (1), 73-91.